ÖREBRO
UNIVERSITET

Lösning till tentamen i

Matematisk statistik och sannolikhetslära

MA506G

2022-01-04

---

1. Vi har en urna med tre sorters kort: två kort som är vita på båda sidor, tre kort med en röd och en vit sida, och två kort som är röda på båda sidor.

   (a) Vi tar upp ett kort på måfå och konstaterar att dess ena sida är röd. Med vilken sannolikhet är även den andra sidan röd? [5p]

   (b) Lägg tillbaka kortet i urnan. Plocka därefter upp två kort utan återläggning och titta på en av deras sidor. Hur stor är sannolikheten att båda korten visar en vit sida? [5p]

   Lösning:

   (a) Det finns 7 röda sidor, varav 4 har en röda omstående sida. Sannolikheten är därmed 4/7.

   Alternativt kan vi låta R_1 vara händelsen att den första sidan är röd och R_2 att omstående sida är röd. Vi har då

$$\mathbf{P}(R_2|R_1) = \frac{\mathbf{P}(R_1 \cap R_2)}{\mathbf{P}(R_1)} = \frac{2/7}{1/2} = \frac{4}{7}.$$

   (b) Låt V_1 vara händelsen att det första kortet visar en vit sida och V_2 att det andra kortet gör det, och låt vidare V_3 vara händelsen det första kortet förutom en synlig vit sida också har en röd sida och V_4 att båda sidorna på det första kortet är vita. Vi har då, genom betingad sannolikhet och lagen om total sannolikhet samt $V_1 = V_3 \cup V_4$, att

$$\begin{aligned}
\mathbf{P}(V_1 \cap V_2) &= \mathbf{P}(V_2|V_1)\mathbf{P}(V_1) \\
&= \mathbf{P}(V_2|V_3)\mathbf{P}(V_3) + \mathbf{P}(V_2|V_4)\mathbf{P}(V_4) \\
&= \frac{6}{12}\frac{3}{14} + \frac{5}{12}\frac{4}{14} \\
&= \frac{18+20}{168} = \frac{19}{84}.
\end{aligned}$$

2. Antag att 48 personer singlar slant 5 gånger var, och låt X beteckna antalet personer som får antingen fem krona eller fem klave.

   (a) Vilken fördelning har X? [3p]

   (b) Beräkna $\mathbf{E}(X)$ och $\mathbf{V}(X)$. [2p]

   (c) Beräkna approximativt $\mathbf{P}(X \leq 3)$. [5p]

   Lösning:

   (a) För varje person gäller att sannolikheten att personen får 5 krona eller 5 klave är 1/16, eftersom sannolikheten att kast 2 till 5 är samma som kast 1 är 1/2 och sannolikheten alla dessa är samma som kast 1 blir då $(1/2)^4$ på grund av oberoende.

   Eftersom vi har 48 personer och respektive person lyckas få 5 lika oberoende av varandra och med samma sannolikhet har vi en binomialbördelning: $X \in \text{Bin}(48, 1/16)$.

   (b) Enligt formelsamlings formler har vi

   $$\mathbf{E}(X) = \frac{48}{16} = 3$$

   och

   $$\mathbf{V}(X) = \frac{48 \cdot 15}{16^2} = \frac{45}{16}.$$

   (c) Vi kan inte approximera med normalfördelning efter $\mathbf{V}(X) < 5$, men däremot med Poissonfördelning eftersom $p < 0{,}1$. Låt $Y \in \text{Po}(3)$. Vi får då, med hjälp av tabell 3,

   $$\mathbf{P}(X \leq 3) \approx \mathbf{P}(Y \leq 3) = 0{,}647.$$

3. Betrakta funktionen

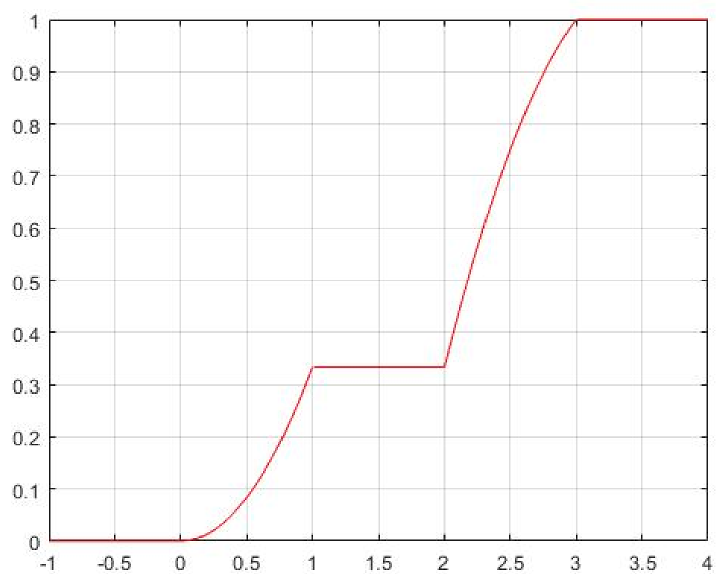
   $$f(x) = \begin{cases} \frac{2}{3}x & 0 \leq x \leq 1 \\ \frac{7}{3} - \frac{2}{3}x & 2 \leq x \leq 3 \\ 0 & \text{annars.} \end{cases}$$

   (a) Beräkna fördelningsfunktionen till X och skissa dess graf. [5p]

   (b) Bestäm väntevärdet för X. [5p]

   Lösning:

(a) Vi har
$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$$
Vi ser direkt att $F(x) = 0$ för $x \leq 0$. För $x \in [0,1]$ gäller
$$F(x) = \int_0^x \frac{2t}{3}\,\mathrm{d}t = \left[\frac{t^2}{3}\right]_0^x = \frac{x^2}{3},$$
och för $x \in [1,2]$ gäller $F(x) = F(1) = 1/3$. Vidare har vi för $x \in [2,3]$ att
$$F(x) = \frac{1}{3} + \int_2^x \left(\frac{7}{3} - \frac{2t}{3}\right)\,\mathrm{d}t = \frac{1}{3} + \left[\frac{7t}{3} - \frac{t^2}{3}\right]_2^x = \frac{7x}{3} - \frac{x^2}{3} - 3,$$
samt $F(x) = F(3) = 1$ för $x > 3$. Fördelningsfunktionens graf ges nedan.



(b) Vi har
$$\begin{aligned}\mathbf{E}(X) = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x &= \int_0^1 x\frac{2x}{3}\,\mathrm{d}x + \int_2^3 x\left(\frac{7}{3} - \frac{2x}{3}\right)\,\mathrm{d}x \\ &= \left[\frac{2x^3}{9}\right]_0^1 + \left[\frac{7x^2}{6} - \frac{2x^3}{9}\right]_2^3 \\ &= \frac{2}{9} - 0 + \frac{21}{2} - 6 - \frac{14}{3} + \frac{16}{9} = \frac{33}{18} = \frac{11}{6}.\end{aligned}$$

4. Låt X vara medeltemperaturen i Örebro under juli. På SMHIs nätsidor kan man läsa att den medeltemperatur som i genomsnitt underskrids vart tionde år är 14,2 °C och den medeltemperatur som i genomsnitt överskrids vart tionde år är 18,0 °C.

   (a) Motivera att X är ungefärligen normalfördelad och beräkna väntevärdet och standardavvikelsen för X? [4p]

   (b) Medeltemperaturen under juli i år var 18,3 °C. Hur stor är sannolikheten att juli under ett godtyckligt valt år är 18,3 °C eller varmare? [3p]

   (c) Vilken värme når man upp till (i genomsnitt) en gång per sekel? (Varmaste juli som man mätt var 1914, med 21,6 °C.) [3p]

   Lösning:

   (a) Den stokastiska variabeln X är medelvärdet av temperaturerna för 31 dagar. Centrala gränsvärdessatsen säger att summan av flera oberoende likafördelade stokastiska variabler går mot en normalfördelning då antalet summander går mot oändligheten, oberoende av summandernas fördelning. Även om dessa temperaturer inte är helt oberoende så kan vi ändå förvänta oss tillräckligt stort oberoende för att påståendet ska gälla. Det är dessutom rimligt att tro att dagsmedeltemperaturerna är normalfördelade, och vi vet att summan av oberoende normalfördelade stokastiska variabler är normalfördelad.

   Enligt tabell 5 gäller $\Phi(0{,}9) = 1{,}2816$ och följdaktligen $\Phi(0{,}1) = -1{,}2816$. Det ger ekvationssystemet

$$\begin{aligned} 14{,}2\ ^\circ\mathrm{C} &= \mu - 1{,}2816; \\ 18{,}0\ ^\circ\mathrm{C} &= \mu + 1{,}2816, \end{aligned}$$

   där μ är väntevärde och σ standardavvikelse för X. Ur dessa två ekvationer löser vi enkelt ut $\mu = 16{,}1\ ^\circ\mathrm{C}$ och $\sigma = 1{,}9\ ^\circ\mathrm{C}/1{,}2816 = 1{,}48\ ^\circ\mathrm{C}$.

   (b) Temperaturen 18,3 °C ligger $2{,}2\ ^\circ\mathrm{C}/1{,}48\ ^\circ\mathrm{C} = 1{,}484$ standardavvikelser över väntevärdet. Sannolikheten att nå så högt är enligt tabell 1

$$\mathbf{P}(X \geq 18{,}3\ ^\circ\mathrm{C}) = 1 - \Phi(1{,}484) = 1 - 0{,}931 = 6{,}9\%.$$

   (c) På nivån 99% ligger vi 2,2363 standardavvikelser över väntevärdet. Det ger temperaturen

$$16{,}1\ ^\circ\mathrm{C} + 2{,}2363 \cdot 1{,}48\ ^\circ\mathrm{C} = 19{,}4\ ^\circ\mathrm{C}.$$

   Det var alltså väldigt varmt 1914.

5. (a) Låt X vara en diskret slumpvariabel med sannolikhetsfunktionen [6p]

$$p_X(k) = (k-1)(1-p)^{k-2}p^2$$

för $k \in \{2, 3, 4, \dots\}$ och $0 \leq p \leq 1$. Vi tar ett stickprov och får då följande värden på X: 16, 8, 9, 10, 27 och 12. Bestäm Maximum-likelihood-skattningen av p.

(b) Låt $Y_j \sim \text{Bin}(5, p)$ för $1 \leq j \leq 2$. Två skattningar av sannolikheten p är [4p]

$$\hat{p}_1 = \frac{Y_1 + Y_2}{10}$$

och

$$\hat{p}_2 = \frac{3Y_1 + Y_2}{20}.$$

Visa att båda skattningarna är väntevärdesriktiga.

Lösning:

(a) Likelihoodfunktionen ges av

$$L(p) = 15(1-p)^{14}p^2 \dots 11(1-p)^{10}p^2 = C(1-p)^{70}p^{12}$$

och dess logaritmerade variant av

$$\ell(p) = \ln\left(C(1-p)^{70}p^{12}\right) = D + 70\ln(1-p) + 12\ln(p).$$

För att hitta maximum deriverar vi och får

$$0 = \ell'(p) = \frac{12}{p} - \frac{70}{1-p},$$

vilket förenklas till ekvationen $70p = 12 - 12p$ med lösningen $p = 6/41$.

(b) Skattningarna är väntevärdesriktiga om deras väntevärden är p. Vi har

$$\begin{aligned}
\mathbf{E}(\hat{p}_1) &= \mathbf{E}\left(\frac{Y_1 + Y_2}{10}\right) \\
&= \frac{1}{10}\left(\mathbf{E}(Y_1) + \mathbf{E}(Y_2)\right) \\
&= \frac{1}{10}(5p + 5p) = p
\end{aligned}$$

och

$$\begin{aligned}\mathbf{E}(\hat{p}_2) &= \mathbf{E}\left(\frac{3Y_1 + Y_2}{20}\right) \\ &= \frac{1}{20}\left(3\mathbf{E}(Y_1) + \mathbf{E}(Y_2)\right) \\ &= \frac{1}{20}\left(15p + 5p\right) = p.\end{aligned}$$

6. Ett tegelbruk vill undersöka om bränntemperatureren påverkar densiteten. Av fem olika lerblandningar skapas två kuber per blandning, där den ena bränns i 150 °C och den andra i 200 °C. Vi får då följande resultat: [10p]

Blandning	1	2	3	4	5
150 °C	33.2	32.9	32.8	32.0	34.5
200 °C	33.6	33.1	33.0	32.9	35.5

Avgör genom att bilda ett 95 % konfidenintervall om de olika förbränningstemperaturerna ger någon skillnad i densitet.

Lösning: Vi har stickprov i par och antar att skillnaderna mellan värderna är normalfördelade enligt $N(\mu, \sigma)$, där μ och σ är okända. Skillnaderna blir, om vi drar bort den kallare från den varmare, $\mathbf{x} = (0.4, 0.2, 0.2, 0.9, 1.0)$. Ur detta material kan vi beräkna medelvädet

$$\bar{x} = \frac{x_1 + \cdots + x_5}{5} = 0.54$$

och stickprovsstandardavvikelsen

$$s = \sqrt{\frac{1}{n-1}\sum_k (x_k - \bar{x})^2} = 0.385.$$

Vi använder sedan intervallskattningen av väntevärdet från formelsamlingen

$$I_\mu = \bar{x} \pm t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}.$$

Med $t_{0.025}(4) = 2.78$ får vi

$$I_\mu = 0.54 \pm 2.78\frac{0.385}{\sqrt{5}} = (0.062, 1.02).$$

Eftersom intervallet inte innehåller 0 så är det skillnad mellan densiteterna för dessa förbränningstemperaturer.